基于攻防博弈和蒙特卡洛模拟的网站防御策略选取方法

吴昊[1]，范九伦[1]，赖成喆[1]，刘建华[2]

（1. 西安邮电大学通信与信息工程学院，陕西 西安 710121；2. 西安邮电大学信息中心，陕西 西安 710121）

摘 要：针对网络攻防对抗中的安全防御策略选取问题，研究攻防双方策略相互影响的动态变化过程。从攻防双方的博弈过程出发，构建攻防博弈模型，基于蒙特卡洛模拟法模拟攻击者的攻击过程，得到攻击者的最佳攻击效用，进而计算防御者的最佳防御效用。该方法实现了在有限的资源投入下选取最优的防御策略，以达到网络安全防御效用的最大化。仿真实验验证了该方法的有效性，并分析了不同参数设置对防御策略选取的影响。

关键词：攻防博弈；防御策略；蒙特卡洛模拟；效用函数

中图分类号：TP393

文献标识码：A

doi: 10.11959/j.issn.1000−436x.2018131

Website defense strategy selection method based on attack-defense game and Monte Carlo simulation

WU Hao[1], FAN Jiulun[1], LAI Chengzhe[1], LIU Jianhua[2]

1. School of Communication and Information Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710121, China

2. Information Centre, Xi'an University of Posts and Telecommunications, Xi'an 710121, China

Abstract: Aiming at the selection of security defense strategy in network attack-defense, the dynamic change process of mutual influence between attack-defense strategy was studied. Based on the game process of both offense and defense, the attack-defense game model was constructed, the attack process of the attacker based on Monte Carlo simulation was simulated and the attacker's best attack utility was obtained, so as to calculate the best defensive utility of the defender. In order to maximize the effectiveness of network security defense, the optimal defense strategy under limited resources was implemented. Simulation experiments verify the effectiveness of the proposed method and analyze the influence of different parameter settings on the selection of defense strategy.

Key words: attack-defense game, defense strategy, Monte Carlo simulation, utility function

1 引言

随着互联网和社会信息化进度的不断发展，网络规模日趋复杂，网络安全问题日益突出。同时，“黑色产业链”带来的每年数千万级的海量病毒和网络攻击，加剧了攻击和防御的不对称，使网络空间安全状况进一步恶化。如果仍然使用传统的被动防御策略已经很难提供有效的防御保障，亟需对网络攻防行为进行分析和预测，并依据分析结果实施有效的主动防御。在安全风险管理中，采用的每种主动防御措施都有一定的防御成本。如何权衡收益和成本是一个具有挑战性的问题，如何在有限的防御成本条件下选取最优的防御策略已成为当前的研究热点[1-2]。

网络攻防的本质就是一种攻防双方策略相互影响的动态变化过程。网络攻防对抗中的攻击方和

收稿日期：2018−03−22；**修回日期**：2018−07−28

基金项目：国家重点研发计划基金资助项目（No.2017YFC0803800）；国家自然科学基金资助项目（No.61671377）；陕西省自然科学基础研究计划基金资助项目（No.2017JQ6010）

Foundation Items: The National Key Research and Development Program of China (No.2017YFC0803800), The National Natural Science Foundation of China (No.61671377), The Natural Science Basic Research Plan of Shaanxi Province (No.2017JQ6010)

防御方是 2 个具有理性思维能力的主体，双方的对立性、策略依存性和关系非合作性正是博弈论的基本特征[3-4]。由于双方在攻防博弈过程中获得的收益有差异，随着时间的推移，在收益差异的牵引和学习机制的驱动下，不断根据对方策略的选择来调整自身策略以确保自身收益，由此形成攻防对抗中的网络安全态势不断动态演化[5-6]。因此，在网络攻防对抗中使用博弈的思想寻找最优的防御策略，是一种很有效的方法[7-10]。文献[11]基于非合作、非零和动态博弈理论提出了完全信息动态博弈主动防御模型，给出了分别适应于完全信息和非完全信息这 2 种场景的攻防博弈算法。文献[12]用三角模糊数来表示攻防双方各策略的损益值，提出了基于三角模糊矩阵的博弈算法。文献[13]基于静态贝叶斯博弈的绩效评估模型 PEM-SBG，提出了蠕虫攻防策略绩效评估方法，通过纳什均衡的计算结果来指导防护策略的选择。文献[14]建立了基于系统动力学的攻防演化博弈模型，通过博弈模型和系统动力方法对攻防双方的策略选取机制进行了分析。文献[15]从动态对抗和有限信息的视角对攻防行为进行研究，构建了攻防信号博弈模型，并提出了精炼贝叶斯均衡求解算法。文献[16]提出了一种基于非合作的演化博弈理论，基于此构建攻防演化博弈模型，提出演化稳定均衡的求解方法。文献[17]设计了多阶段攻防博弈均衡的求解方法，并给出了最优主动防御策略选取算法。文献[18]结合进化博弈论和马尔可夫决策过程，构造了一个基于有限理性约束的网络攻防分析的多级马尔可夫进化博弈模型，能够对多阶段、多状态网络攻防过程进行动态分析和推理。文献[19]基于不完全信息动态博弈构建了面向动态目标防御的单阶段和多阶段博弈模型，给出了精炼贝叶斯均衡求解算法和先验信息修正方法，获得了不同安全态势下的最优动态目标防御策略。文献[20]借鉴传染病动力学理论构建了攻防微分博弈模型，并提出了鞍点策略的求解方法和最优防御策略的选取算法。文献[21]设计了一种基于无标度多目标零和博弈的网络攻击建模方法，使用帕累托前沿确定最有害的攻击，并利用帕累托优化来找到对这些攻击的最佳防御。但这些方案大都没考虑到网络安全防御资源投入受限的情况。在实际的应用过程中，尤其是一些中小企业的网络系统，在网络安全防御资源的投入上往往会受到人力、物力、财力等各方面的限制，很可能无法达到理论上最优防御策略的要求。因此，如何在防御资源投入总额限定的情况下做出最优的防御决策，实现网络安全防御效用的最大化，是本文着力解决的问题。

网络安全防御应平衡网络安全与系统资源投入的关系，采用“适度安全”的攻防策略，正是基于这一目标，本文在研究攻防博弈的基础上，结合蒙特卡洛模拟法模拟攻击者的攻击，得到攻击者的最佳攻击效用，进而使防御者在有限可选策略集合中选取最优的防御策略。随后通过实验对该方法的有效性进行了验证，实验结果表明，该方法能够通过完整模拟攻防过程中攻防双方的博弈过程，帮助管理员在防御资源投入总额受限的情况下选取最优的防御资源配置，为安全防御决策提供实用而有效的指导。

2　基于攻防博弈和蒙特卡洛模拟的网站防御策略

2.1　网络攻防博弈模型

假设某站点可能受到攻击群体 A 的攻击，防御者 D 需要部署防御资源 $d \in D$ 来保护站点防御攻击者的攻击。但由于经济条件所限，只能根据攻击情况调整自己的防御策略，在有限的防御资源中选择最优的配置资源。攻击者则通过观察防御者的决策 d，计算得到合适的攻击决策 $a \in A$，以选择合适的攻击方法，如不同的攻击工具、攻击时间间隔、攻击者人数等。由于防御者的防御决策 d 和攻击者的攻击决策 a 之间的相互影响，最终导致一个随机结果 $s \in S$。因此，防御者可以根据不同的防御决策 d 和结果 s，得到一个防御序列 c_D，进而得到防御效用 u_D。攻击者同样可以根据攻击决策 a 和结果 s，得到攻击序列 c_A，继而得到攻击效用 u_A。整个攻防博弈过程如图 1 所示。

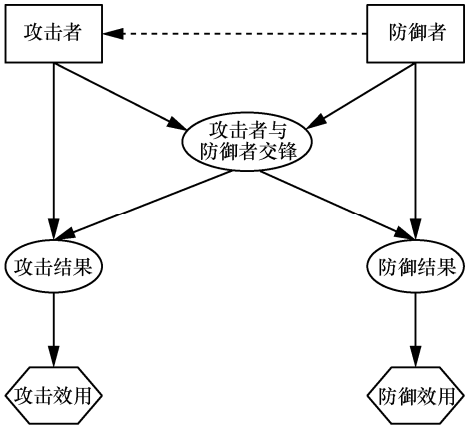

图 1　攻防博弈过程

2.2 防御者的防御目标

防御者的目标就是通过计算防御效用 $u_D(d,s)$ 来评估不同的防御结果，进而找到最佳的防御策略 d^*。具体方法如下。

对于给定的防御资源 d 和攻击决策 a，首先需要建立概率模型 $p_D(s|d,a)$（表示当防御者部署了防御资源 d 以后，攻击者继续发动攻击 a 并产生结果 s 的可能性），通过计算此时的预期防御效用 $u_D(d,s)$ 之后，就可以得到给定攻击的防御效用，为

$$\psi_D(d|a)=\int u_D(d,s)p_D(s|d,a)\mathrm{d}s \tag{1}$$

这时，如果防御者能够建立攻击概率模型 $p_D(a|d)$（表示当攻击者发现了防御决策 d 后，最有可能采用的攻击决策 a 的概率分布），就可以得到防御者的总体效用函数，为

$$\Psi_D(d)=\int \psi_D(d|a)p_D(a|d)\mathrm{d}a \tag{2}$$

此时，防御者只要找到 $\max\limits_d \Psi_D(d)$ 时 d 的取值 d^*，即为防御者的最佳防御资源决策。

2.3 攻击概率模型估计

在式(2)的计算中，由于防御者无法直接得到攻击者的攻击决策，也就无法直接得到攻击概率模型 $p_D(a|d)$，因此只能通过模仿攻击者，然后对每种攻击决策分别估算相应的概率。具体原理如下。

攻击者为了达到最佳的攻击效用，也需要得到相关的攻击效用 $u_A(a,s)$ 和攻击概率 $p_A(s|d,a)$（具体计算方法见第 3 节），通过式(3)找到已知防御 d 后的最佳攻击 a^*。

$$a^*(d)=\arg\max_{a\in A}\int u_A(a,s)p_A(s|d,a)\mathrm{d}s \tag{3}$$

然而，防御者很难得到攻击者的 u_A 和 p_A。因此我们采用蒙特卡洛模拟法[21-23]，进行一系列的随机攻击模拟实验，得到已知防御 d 时相应的攻击概率，进而建立攻击模型 (U_A,P_A)，由式(3)可得

$$A^*(d)=\arg\max_{a\in A}\int U_A(a,s)P(s|d,a)\mathrm{d}s \tag{4}$$

这样，我们就可以将所有实验中攻击效用最大时的攻击概率作为攻击者 A 的最优攻击概率，此时的攻击决策 a 为攻击者的最优攻击决策，即 $p_D(a|d)=P(A^*(d)\leqslant a)$。具体过程如算法 1 所示。

算法 1 最优攻击决策算法

begin

输入模拟攻击实验次数 K

for $k=1$ to K

随机生成 $a_k\in A$

计算 $u_A(a_k,s)$， $p_A(s|d,a_k)$

$\psi_A(d,a_k)\sim\int u_A(a_k,s)p_A(s|d,a_k)\mathrm{d}s$

$a^*(d)\sim\arg\max\limits_{a_k\in A}\psi_A^k(d,a_k)$

$p_D(a|d)=P(A^*(d)\leqslant a)$

end

由算法 1 得到攻击概率模型 $p_D(a|d)$ 后，我们就可以利用式(2)找到防御者的最佳防御资源决策。

3 最优防御策略选取实例

为了后续计算方便，假设某网站可选的防御资源共 5 类，记作 $d_p=(d_1,d_2,d_3,d_4,d_5)$，分别代表服务器硬件设计、服务器软件升级、防火墙设备、网站管理人员和安全审计系统的数量，m、n、j、k、l 分别为 d_1、d_2、d_3、d_4、d_5 可取的最大值，q_1、q_2、q_3、q_4、q_5 分别为每种资源的单价。每种资源的功用和相关系数如表 1 所示。其中，$l=1$，即 $d_5\in\{0,1\}$，$d_5=1$ 表示安全审计系统，$d_5=0$ 表示没有安全审计系统。管理员需要根据不同的网络攻击危险程度，调整防御资源配置，并根据不同的防御配置和防御效果计算相应的防御效用。

管理员的防御投资为

$$C_{\mathrm{inv}}(d_1,d_2,d_3,d_4,d_5)=q_1d_1+q_2d_2+q_3d_3+q_4d_4+q_5d_5 \tag{5}$$

假设防御的总投资不能超过 B，则防御资源配

表 1 每种资源的功用和相关系数

防御资源	资源功用	相关系数
服务器硬件设计	提高网站服务器的工作能力，降低被攻击的可能性	u_1
服务器软件升级	修补软件漏洞，防止外部攻击	u_2
防火墙设备	对网站出入流量进行管理，阻止攻击流量的进入	u_3
网站管理人员	对网站安全配置进行管理，同时检查系统日志等信息，从而发现被攻击的线索	u_4、ρ_4
安全审计系统	不直接对网络攻击进行防御，但当系统被攻击后，可以及时发现并保存攻击证据	ρ_5

置 $d_p=(d_1,d_2,d_3,d_4,d_5)$ 应满足以下条件。

1) $q_1d_1+q_2d_2+q_3d_3+q_4d_4+q_5d_5\leqslant B$ 。

2) $d_1,d_2,d_3,d_4,d_5\geqslant 0$，且 d_1、d_2、d_3、d_4、d_5 均为整数。

3) $d_1\leqslant m$，$d_2\leqslant n$，$d_3\leqslant j$，$d_4\leqslant k$，m、n、j、k 分别为 d_1、d_2、d_3、d_4 可取的最大值。

4) $d_5\in\{0,1\}$。

3.1　攻击者决策过程和攻击效用计算

网络攻击者 A 通过观察防御者的防御配置来制定相应的攻击决策。具体过程如图 2 所示。

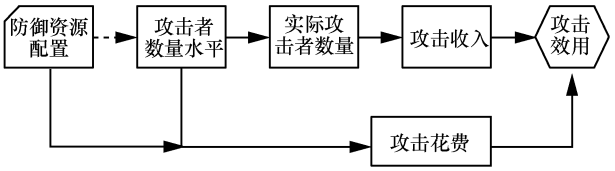

图 2　攻击者决策过程

1) 攻击者看到了管理员的防御投资 $d_p=(d_1,d_2,d_3,d_4,d_5)$。

2) 决定攻击者的数量 $t\in A$， A 为所有网络攻击者的集合。

3) 实际的攻击者数量 $t'=(1-\tau(d_p))t$，因为实际的攻击过程中一些攻击者可能提前放弃攻击。$\tau(d_p)$ 为看到防御资源 d_p 时提前放弃的攻击者的比率。假设 τ 服从贝塔分布[24]，即 $\tau(d_p)\sim \mathrm{B}(\alpha_\tau(d_p),\beta_\tau(d_p))$，如果所有的攻击者都可能完成攻击，$\tau$ 接近于 0，因此 $\alpha_\tau\ll\beta_\tau$。此时，可得到 $p_A(\tau(d_p))$ 是一个狄利克雷过程[25-26]，基分布为 $p_D(\tau(d_p)$，集中参数为 δ_τ，即 $p_A(\tau(d_p))\sim \mathrm{DP}(p_D(\tau(d_p),\delta_\tau)$，由随机分布 $p_A(\tau(d_p))=p_A(\tau\,|\,d_p)$ 可得到 $p_A(t'\,|\,t,d_p)$。

4) 攻击者实施攻击的准备成本是 q_pt，q_p 为每个攻击者的平均准备成本，包括购买攻击工具成本、准备时间成本等。

5) 每次攻击都面临的攻击成本如下。

①假设攻击无法进行的概率为 $1-\xi$，此时唯一的损失是准备成本，攻击成功进行的概率 ξ 依赖于攻击者的攻击水平。同时，攻击能否进行也受防御资源 (d_1,d_2,d_3,d_4) 的影响，但随着防御资源数量的增加，对攻击成功概率的影响会逐渐降低。假设 ξ 也服从贝塔分布，即 $\xi(d_1,d_2,d_3,d_4)\sim \mathrm{B}(\alpha_\xi(d_1,d_2,d_3,d_4),\beta_\xi(d_1,d_2,d_3,d_4))$，其方差为 σ_ξ，均值为

$$\mathrm{E}(\xi\,|\,d_1,d_2,d_3,d_4)=\xi_0\exp(-\mu_1d_1-\mu_2d_2-\mu_3d_3-\mu_4d_4)+\xi_r \tag{6}$$

其中，μ_1、μ_2、μ_3、μ_4 为防御资源 (d_1,d_2,d_3,d_4) 每增加一个单位，攻击成功率减少的比率。$\xi_0+\xi_r$ 为目前没有额外防御资源增加情况下的攻击成功概率，ξ_r 为即使资源 (d_1,d_2,d_3,d_4) 无限量部署，攻击也会进行的概率。此时，$P_A(\xi(d_1,d_2,d_3,d_4))$ 也是一个狄利克雷过程，即 $P_A(\xi(d_1,d_2,d_3,d_4))\sim \mathrm{DP}(P_D(\xi(d_1,d_2,d_3,d_4)),\delta_\xi)$。

②攻击者实施了攻击但被抓到的概率为 $\xi\theta$，每个攻击者面临的平均犯罪罚款为 f_p，被抓的概率 θ 取决于安全审计系统和网站管理人员的能力和数量。但这种依赖是非线性的，随着安全审计系统和网站管理人员数量的增加，对抓住攻击者概率的影响会逐渐变小。假设 θ 也服从贝塔分布，即 $\theta(d_4,d_5)\sim \mathrm{B}(\alpha_\theta(d_4,d_5),\beta_\theta(d_4,d_5))$，其方差为 σ_θ，均值为

$$\mathrm{E}(\theta\,|\,d_4,d_5)=1-\exp(-\rho_4d_4-\rho_5d_5) \tag{7}$$

其中，(ρ_4,ρ_5) 为防御资源 (d_4,d_5) 每增加一个单位，攻击者被抓的可能性增加的比率。此时，$P_A(\theta(d_4,d_5))$ 也是一个狄利克雷过程，即 $P_A(\theta(d_4,d_5))\sim \mathrm{DP}(P_D(\theta(d_4,d_5)),\delta_\theta)$。

③攻击者成功实施攻击且不被抓住的概率为 $\xi(1-\theta)$，每个攻击成功者所能得到的收益为 ω，假设 ω 服从均匀分布[27]，即 $\omega\sim \mathrm{U}(\omega_a,\omega_b)$，$\omega_a$、$\omega_b$ 分别为收益的最低值和最高值。

6) 此时，攻击者预期的收益为 $c_A(t_1,t_2,t_3)=\omega t_3-q_pt-f_pt_2$，其中，$(t_1,t_2,t_3)$ 服从多项式分布 $M(t';1-(1-\tau)\xi,(1-\tau)\xi\theta,(1-\tau)\xi(1-\theta))$。

7) 攻击者实施攻击获得的收益与其付出的风险成本成正比。风险越大，收益可能越大。因此，我们将攻击者的攻击效用函数定义为

$$u_A(c_A)=\exp(k_Ac_A),k_A>0 \tag{8}$$

其中，k_A 为攻击者风险系数。由于 k_A 随攻击者的攻击状态不断变化，本文假设 k_A 的最大值为 K_A，且服从均匀分布，即 $k_A\sim \mathrm{U}(0,K_A)$，则攻击者的随机攻击效用为

$$u_A(c_A)=\exp(k_Ac_A),k_A\sim \mathrm{U}(0,K_A) \tag{9}$$

令 $p_{t_1t_2t_3d_p}=\Pr(t_i|d_p,i=1,2,3)$，$d_p$ 表示防御投资，则有

$$P_A(\xi(d_1,d_2,d_3,d_4))=P_A(\xi\mid d_1,d_2,d_3,d_4)$$
$$P_A(\theta(d_4,d_5))=P_A(\theta\mid d_4,d_5)$$

可以得到攻击者的随机期望效用为

$$\Psi_A(t',t,d_p)=\iint[\sum_{t_1,t_2,t_3}p_{t_1t_2t_3d_p}U_A(\omega t_3-q_pt-f_pt_2)]\cdot P_A(\xi\mid d_1,d_2,d_3,d_4)P_A(\theta\mid d_4,d_5)\mathrm{d}\xi\mathrm{d}\theta$$

根据 t' 的不确定性，得到随机期望函数为

$$\Psi_A(t,d_p)=\int\Psi_A(t',t,d_p)P_A(t'\mid t,d_p)\mathrm{d}t' \quad (10)$$

这样，我们只要通过算法 1 找到防御资源为 d_p 时，攻击者数量的最优值 $T^*(d_p)=\arg\max_{t\in A}\Psi_A(t,d_p)$，即为最优的攻击者数量水平。同理可得

$$\int p_D(t\mid d_p)\mathrm{d}t=p_D(T\leqslant t\mid d_p)=P(T\leqslant t\mid d_p)=P(T^*(d_p)\leqslant t) \quad (11)$$

3.2 防御者防御效用计算

假设现有防御资源的防御能力为 b，防御者根据攻击者的数量改变现有的防御资源配置，从而改变现有的防御能力。防御资源的多少直接影响防御能力的高低，同时，当防御资源一定时，随着网站周围攻击者数量 t 的增加，网站的防御能力 b 会不断降低。因此，我们使用逻辑回归模型[28]描述攻击者数量与防御能力之间的关系，如图 3 所示。其中，横轴代表攻击者数量 t，纵轴代表防御能力 b。

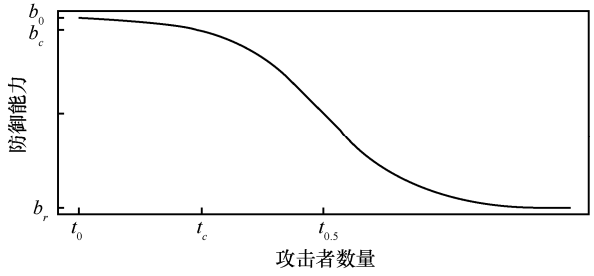

图 3 攻击者数量与防御能力关系

由于 $b\mid t$ 具有不确定性，假设 $p_D(b\mid t)$ 服从正态分布，因此可以通过 t 的分布得到 b 的分布。其方差为 δ_b，均值为

$$\mathrm{E}[b\mid t]=\frac{b_0-b_r}{1+\exp[\gamma_b(t-t_{0.5})]}+b_r,t>0 \quad (12)$$

其中，$b=b_0$ 为理想的防御能力，此时网站周围没有潜在的攻击者，完全不存在被攻击的危险（当 $t=t_0\ll t_{0.5}$ 时，我们即认为当前攻击者的数量接近于 0），当前周围攻击者数量为 $t=t_c$ 时对应的防御能力为 $b=b_c$，如果攻击者数量达到最大值，此时的防御能力下降到最小值，即 $b=b_r$。当攻击者数量超过一定的阈值时，我们就认为防御能力 b 无限接近 b_r。$t=t_{0.5}$ 表示当防御能力下降一半，即 $b=\frac{b_r-b_0}{2}$ 时的攻击者数量。γ_b 表示随着攻击者数量的增加，防御能力下降的速度。

由于防御能力 b 的变化可以直接反映防御资源的价值，因此用防御能力 b 直接表示防御资源的价值，可得到管理员的收益为

$$C_D(d_1,d_2,d_3,d_4,d_5,b)=-C_{\mathrm{inv}}(d_1,d_2,d_3,d_4,d_5)-(b_0-b) \quad (13)$$

管理员希望不断规避风险，同时尽可能少地投入，本文将管理员的防御效用定义为

$$u_D(c_D)=\exp(-k_Dc_D),\quad k_D>0 \quad (14)$$

其中，k_D 为防御者风险系数，则防御者的总体防御效用函数为

$$\Psi_D(d)=\iint u_D(c_D)p_D(t\mid d_p)p_D(b\mid t)\mathrm{d}t\mathrm{d}b \quad (15)$$

由于 $P(T(d_p)\leqslant t)=\int p_D(t\mid d_p)\mathrm{d}t$，$P(b_c\leqslant b\mid t)=\int p_D(b\mid t)\mathrm{d}b$，由式(11)和式(15)可得

$$\Psi_D(d)=u_D(c_D)P(T^*(d_p)\leqslant t)P(b_c\leqslant b\mid t=T^*(d_p)) \quad (16)$$

这样，我们只要根据 $\arg\max\Psi_D(d)$ 找到相应的 d^*，即为最优防御配置。

4 策略选取结果分析

4.1 参数选取

假设某网站每年的安全防御预算不超过 30 万元，每种资源的单价和最大数量如表 2 所示。

表 2 防御资源单价及最大数量

防御资源	单价/万元	最大数量
服务器硬件升级	2	3
服务器软件升级	1.5	3
防火墙设备	5	3
网站管理人员	10（管理人员每年工资）	2
安全审计系统	7	1

模型中其他参数选取如下。

1) 假设目前试图对该网站进行攻击的人数 $t=t_c$ 服从二项分布，均值为 100，方差为 25，即 $t_c\sim\mathrm{Bin}(200,0.5)$。

2) 提前放弃攻击的攻击者比率 $\tau \sim \mathrm{B}(1,9)$。

3) 系统现有理想的防御能力 $b_0 = 35$ 万元，即相当于防御资产价值为 35 万元时，系统周围没有潜在攻击者时的防御能力，现有的实际防御能力为 $b_c = 30$ 万元，防御能力最低为现有防御能力的 80%，即 $b_r = 24$ 万元。

4) 假设当 $t_{0.5} = 100$ 时，防御能力下降一半。同时取 $\gamma_b = 0.08$，$\sigma_b = 3$，管理员可以根据式(9)估算出 $\mathrm{E}[b \mid t]$。

5) 攻击者的准备成本 $q_p = 0.1$ 万元，平均犯罪罚款 $f_p = 2$ 万元。

6) 假设攻击者成功实施攻击，他们每个人所获得的收益 $\omega \sim \mathrm{U}(1,2)$ 万元。

7) 假设 $\xi_0 + \xi_r \sim B(3,1)$，$\xi_r = 0.05$，即有 5%的攻击者无论如何都要实施攻击。同时取 $u_1 = 0.2$，$u_2 = 0.15$，$u_3 = 0.7$，$u_4 = 0.25$。取 $\delta_\xi = 0.05$，可以根据式(6)估算出 $\mathrm{E}(\xi \mid d_1, d_2, d_3, d_4)$。

8)取 $\rho_4 = 0.25$，$\rho_5 = 0.5$，$\delta_\theta = 0.08$，可以根据式(7)估算出 $\mathrm{E}(\theta \mid d_4, d_5)$。

9)取防御者风险系数 $k_D = 0.05$，攻击者风险系数 $k_A \sim \mathrm{U}(0,0.1)$。

4.2 实验结果及分析

由于防御总预算不超过 30 万元，可行的防御方案总共有 241 种，每种方案测试 50 次，图 4 显示了每种方案的平均防御效用。我们选取其中最优的 5 种方案，它们的投资和防御效用如表 3 所示。

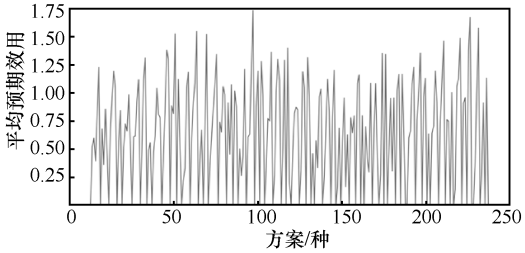

图 4　每种方案的平均防御效用

表 3　最优防御方案 Top5

排名	防御策略	总投入/万元	平均预期效用
1	(1,1,1,2,0)	28.5	1.759
2	(3,2,2,1,0)	29.0	1.698
3	(3,3,0,1,1)	27.5	1.599
4	(0,3,1,2,0)	29.5	1.571
5	(0,2,2,1,1)	30.0	1.548
平均		28.9	1.635

由表 3 可知，排名第一的防御策略的投入比第二名、第四名、第五名的投入都低，这说明好的防御策略并不是投入得越多越好，而是根据周边的威胁情况，选择合适的防御方案。这样不仅投入少，还能获得更好的防御效果。

4.3 参数选取对防御策略选取的影响

为了充分了解算法中不同的参数与防御策略的选取之间的内在关系，下面分别从提高犯罪罚款和提高攻击收益这 2 个方面分析不同的参数选取对防御策略的影响。

1) 提高犯罪罚款

假设攻击者被抓后的平均罚款增加 50%，即 $f_p = 3$ 万元，其他参数不变，重新进行实验，得到的结果如图 5 所示，排名前五的防御策略如表 4 所示。

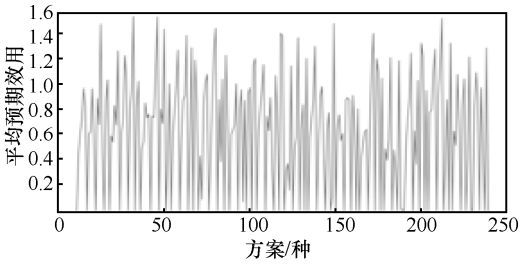

图 5　提高攻击者犯罪罚款后每种方案的平均防御效用

表 4　提高攻击者犯罪罚款后的最优防御方案 Top5

排名	防御策略	总投入/万元	平均预期效用
1	(0,1,2,1,1)	28.5	1.591
2	(0,2,1,2,0)	28.0	1.587
3	(3,1,1,1,1)	29.5	1.577
4	(2,0,3,1,0)	29.0	1.534
5	(0,0,2,1,1)	27.0	1.528
平均		28.4	1.563

由表 4 可知，前五名防御策略的平均投入为 28.4 万元，相较原先的平均投入 28.9 万元有一定的减少，由此可以看出，当提高犯罪罚款标准以后，网络攻击者由于忌惮被抓后高额的罚款，会减少攻击，这样防御者可以用更少的投资即可达到预期的防御效果。由此可见，只有加大对危害计算机信息系统安全犯罪的打击力度，才能有效地震慑计算机犯罪，保障计算机信息系统安全和信息安全，促进我国互联网的健康发展。

2) 提高攻击收益

假设攻击者攻击成功后的平均收益提高到

$\omega \sim \mathrm{U}(2,3)$，其他参数不变，重新进行实验，得到的结果如图 6 所示，排名前五的防御策略如表 5 所示。

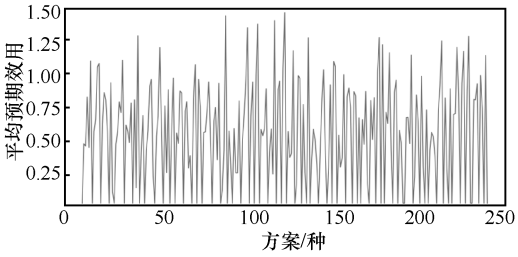

图 6　提高攻击平均收益后每种方案的平均防御效用

表 5　提高攻击平均收益后的最优防御方案 Top5

排名	防御策略	总花费/万元	平均预期效用
1	(1,2,3,1,0)	30.0	1.656
2	(1,0,2,1,1)	29.0	1.630
3	(1,2,1,2,0)	30.0	1.588
4	(1,1,3,1,0)	28.5	1.558
5	(1,1,1,2,0)	28.5	1.527
平均		29.2	1.592

由表 5 可知，前五名防御策略的平均投入为 29.2 万元，相较原先的平均投入 28.9 万元进一步提高，其原因在于随着攻击预期收益的增加，攻击者会更加铤而走险地发动攻击，防御者就得相应地增加防御投资来抵御被攻击的风险。因此，对一些保存有重要信息和资源的网站，更应该增加网络安全防御的投入，以加强对周边安全威胁的主动防御。

5　结束语

通过攻防博弈模型对网络攻防行为进行分析和预测，可以帮助系统管理员及时发现网络中潜在的网络威胁和存在的安全隐患[29]，根据不同的安全需求进行最优防御策略的选择和实施，最大效率地达到网络安全防护的目的。

本文从攻防双方的博弈过程出发，构建攻防博弈模型，结合蒙特卡洛模拟法模拟攻击者的攻击，从而得到攻击者的最佳攻击效用，进而获得防御者在防御资源投资总额限定下的最优防御策略。实验结果表明，该方法有效可行，可以帮助网站管理人员通过预先估计攻击者各种攻击行为可能性的变化，采取更高效的防御手段，提升防御效果。

参考文献：

[1] ZHANG H G, HAN W B, LAI X J, et al. Survey on cyberspace security[J]. Science China Information Sciences, 2015, 58(11): 1-43.

[2] 龚俭, 臧小东, 苏琪, 等. 网络安全态势感知综述[J]. 软件学报, 2017, 28(4): 1010-1026.
GONG J, ZANG X D, SU Q, et al. Survey of network security situation awareness[J]. Journal of Software, 2017, 28(4): 1010-1026.

[3] 姜伟, 方滨兴, 田志宏, 等. 基于攻防随机博弈模型的防御策略选取研究[J]. 计算机研究与发展, 2010, 47(10): 1714-1723.
JIANG W, FANG B X, TIAN Z H, et al. Research on defense strategies selection based on attack-defense stochastic game model[J]. Journal of Computer Research & Development, 2010, 47(10): 1714-1723.

[4] 王元卓, 于建业, 邱雯, 等. 网络群体行为的演化博弈模型与分析方法[J]. 计算机学报, 2015, 38(2): 282-300.
WANG Y Z, YU J Y, WEN Q, et al. Evolutionary game model and analysis methods for network group behavior[J]. Chinese Journal of Computers, 2015, 38(2): 282-300.

[5] LIANG X, XIAO Y. Game theory for network security[J]. IEEE Communications Surveys & Tutorials, 2013, 15(1): 472-486.

[6] ROY S, ELLIS C, SHIVA S, et al. A survey of game theory as applied to network security[C]//Hawaii International Conference on System Sciences. 2010: 1-10.

[7] YANG R, KIEKINTVELD C, ORDONEZ F, et al. Improving resource allocation strategies against human adversaries in security games: an extended study[J]. Artificial Intelligence, 2013, 195(1): 440-469.

[8] FALLAH M. A puzzle-based defense strategy against flooding attacks using game theory[J]. IEEE Transactions on Dependable & Secure Computing, 2010, 7(1): 5-19.

[9] CHENG D, HE F, QI H, et al. Modeling, analysis and control of networked evolutionary games[J]. IEEE Transactions on Automatic Control, 2015, 60(9): 2402-2415.

[10] 王元卓, 林闯, 程学旗, 等. 基于随机博弈模型的网络攻防量化分析方法[J]. 计算机学报, 2010, 33(9): 1748-1762.
WANG Y Z, LIN C, CHENG X Q, et al. Analysis for network attack-defense based on stochastic game model[J]. Chinese Journal of Computers, 2010, 33(9): 1748-1762.

[11] 林旺群, 王慧, 刘家红, 等. 基于非合作动态博弈的网络安全主动防御技术研究[J]. 计算机研究与发展, 2011, 48(2): 306-316.
LIN W Q, WANG H, LIU J H, et al. Research on active defense technology in network security based on non-cooperative dynamic game theory[J]. Journal of Computer Research & Development, 2011, 48(2): 306-316.

[12] 高翔, 祝跃飞, 刘胜利. 应用三角模糊矩阵博弈的网络安全评估研究[J]. 西安交通大学学报, 2013, 47(8): 49-53.
GAO X, ZHU Y F, LIU S L. Networks security assessment based on triangular fuzzy matrix game[J]. Journal of Xi'an Jiaotong University, 2013, 47(8): 49-53.

[13] 刘玉岭, 冯登国, 吴丽辉, 等. 基于静态贝叶斯博弈的蠕虫攻防策略绩效评估[J]. 软件学报, 2012, 23(3): 712-723.
LIU Y L, FENG D G, WU L H, et al. Performance evaluation of worm attack and defense strategies based on static Bayesian game[J]. Journal of Software, 2012, 23(3): 712-723.

[14] 朱建明, 宋彪, 黄启发. 基于系统动力学的网络安全攻防演化博弈模型[J]. 通信学报, 2014, 35(1): 54-61.
ZHU J M, SONG B, HUANG Q F. Evolution game model of offense-defense for network security based on system dynamics[J].

Journal on Communications, 2014, 35(1): 54-61.

[15] 张恒巍, 余定坤, 韩继红, 等. 基于攻防信号博弈模型的防御策略选取方法[J]. 通信学报, 2016, 37(5): 51-61.
ZHANG H W, YU D K, HAN J H, et al. Defense policies selection method based on attack-defense signaling game model[J]. Journal on Communications, 2016, 37(5): 51-61.

[16] 黄健明, 张恒巍, 王晋东, 等. 基于攻防演化博弈模型的防御策略选取方法[J]. 通信学报, 2017, 38(1): 168-176.
HUANG J M, ZHANG H W, WANG J D, et al. Defense strategies selection based on attack-defense evolutionary game model[J]. Journal on Communications, 2017, 38(1): 168-176.

[17] 张恒巍, 李涛. 基于多阶段攻防信号博弈的最优主动防御[J]. 电子学报, 2017, 45(2): 431-439.
ZHANG H W, LI T. Optimal active defense based on multi-stage attack-defense signaling game[J]. Acta Electronica Sinica, 2017, 45(2): 431-439.

[18] HUANG J, ZHANG H, WANG J. Markov evolutionary games for network defense strategy selection[J]. IEEE Access, 2017, PP(99): 1.

[19] 刘江, 张红旗, 刘艺. 基于不完全信息动态博弈的动态目标防御最优策略选取研究[J]. 电子学报, 2018, 46(1): 82-89.
LIU J, ZHANG H Q, LIU Y. Research on optimal selection of moving target defense policy based on dynamic game with incomplete information[J]. Acta Electronica Sinica, 2018, 46(1): 82-89.

[20] 张恒巍, 李涛, 黄世锐. 基于攻防微分博弈的网络安全防御决策方法[J]. 电子学报, 2018, 46(6): 1428-1435.
ZHANG H W, LI T, HUANG S R. Network defense decision-making method based on attack-defense differential game[J]. Acta Electronica Sinica, 2018, 46(6): 1428-1435.

[21] SUN Y, XIONG W, YAO Z, et al. Analysis of network attack and defense strategies based on pareto optimum[J]. Electronics, 2018, 7(3): 36.

[22] SEILA A. Simulation and the Monte Carlo method[J]. Technometrics, 2009, 24(2): 167-168.

[23] RUBINSTEIN R Y, KROESE D P. Simulation and the Monte Carlo method, second edition[M]. Wiley New York, 2007.

[24] 林要华, 梁忠, 胡华平. 贝塔分布的布谷鸟搜索算法[J]. 南京大学学报, 2016, 52(4): 638-646.
LIN Y H, LIANG Z, HU H P. Cuckoo search algorithm with beta distribution[J]. Journal of Nanjing University, 2016, 52(4): 638-646.

[25] 梅素玉, 王飞, 周水庚. 狄利克雷过程混合模型、扩展模型及应用[J]. 科学通报, 2012, 57(34): 3243-3257.
MEI S Y, WANG F, ZHOU S G. Dirichlet process mixture model, extensions and applications[J]. Chinese Journal, 2012, 57(34): 3243-3257.

[26] 严宇宇, 陶煜波, 林海. 基于层次狄利克雷过程的交互式主题建模[J]. 软件学报, 2016, 27(5): 1114-1126.
YAN Y Y, TAO Y B, LIN H. Interactive topic modeling based on hierarchical Dirichlet process[J]. Journal of Software, 2016, 27(5): 1114-1126.

[27] 常诗雨, 宋礼鹏. 基于演化博弈论的网络安全投资策略分析[J]. 计算机工程与设计, 2017, 38(3): 611-615.
CHANG S Y, SONG L P. Analysis of network security investment strategy based on evolutionary game theory[J]. Computer Engineering & Design, 2017, 38(3): 611-615.

[28] ZHANG S, ZHANG L, QIU K, et al. Variable selection in logistic regression model[J]. Chinese Journal of Electronics, 2015, 24(4): 813-817.

[29] 周靖哲, 陈长松. 云计算架构的网络信息安全对策分析[J]. 信息网络安全, 2017(11): 74-79.
ZHOU J Z, CHEN C S. Analysis of network information security in the cloud computing architecture[J]. Netinfo Security, 2017(11): 74-79.

[作者简介]

吴昊（1981−），男，江苏武进人，西安邮电大学讲师，主要研究方向为信息安全。

范九伦（1964−），男，河南温县人，博士，西安邮电大学教授，主要研究方向为信号处理和信息安全。

赖成喆（1985−），男，陕西汉中人，博士，西安邮电大学副教授，主要研究方向为信息安全。

刘建华（1963−），男，陕西宝鸡人，西安邮电大学高级工程师，主要研究方向为信息安全。